

E-Commerce Churn Analysis: Data-Driven Insights to Strengthen Retention Strategy at Shopshop



Risman Ariska Sultani

TABLE OF CONTENT

Company Background

Business Context & Stakeholder
Involved

Data Understanding

Exploratory Data Analysis

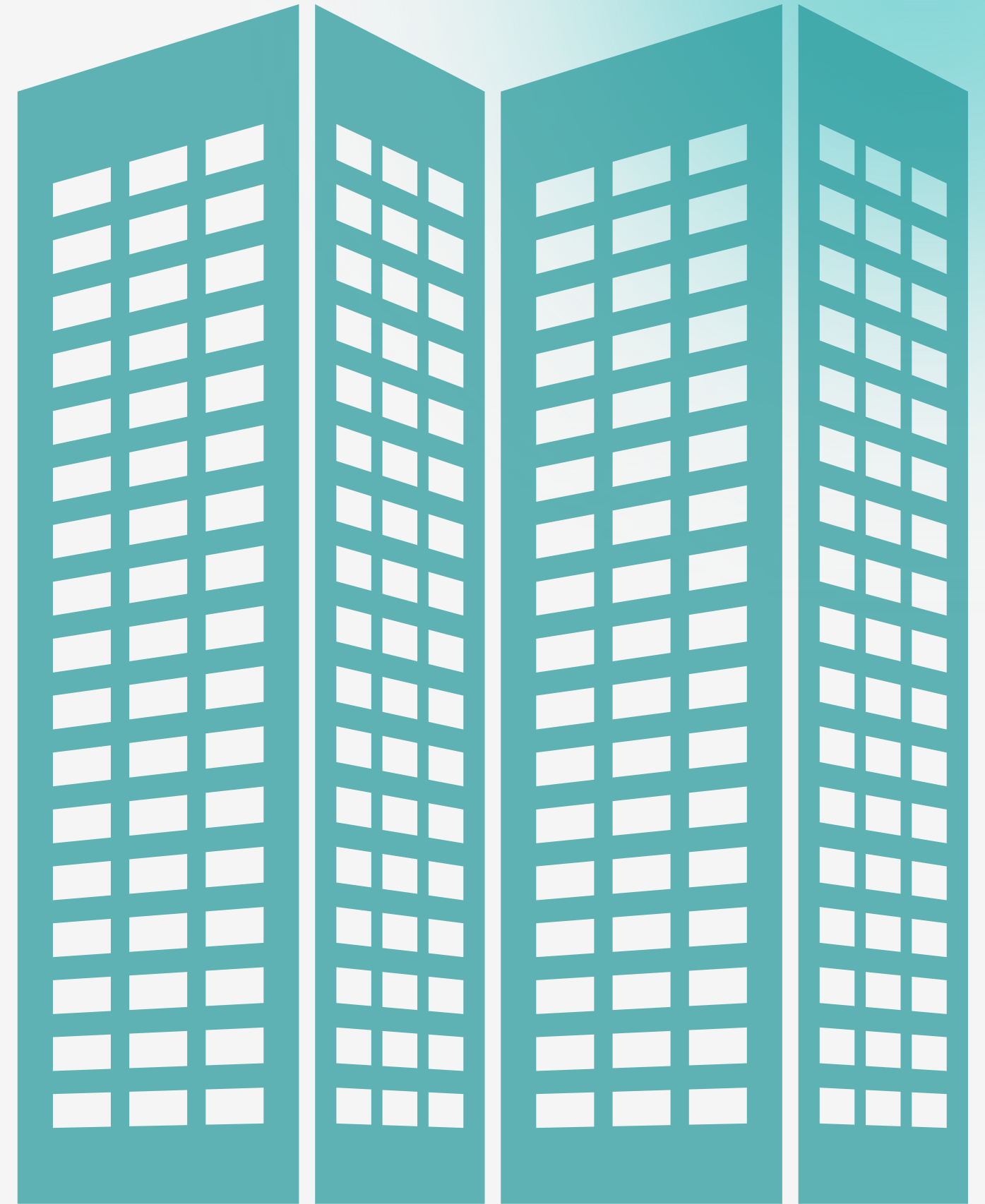
Conclusion & Recommendation

COMPANY BACKGROUND

About Our Company

Shopshop is a national-scale e-commerce company that offers a wide range of everyday products through its mobile application and website.

The company serves thousands of active customers with diverse profiles — varying in age, location, device preferences, payment methods, and shopping behavior.



BUSINESS CONTEXT & PROBLEM STATEMENT

STAKEHOLDER INVOLVED



BUSINESS CONTEXT

In the e-commerce industry, competition is extremely intense. The cost of acquiring new customers is often much higher than the cost of retaining existing ones.

Recently, Shopshop has also experienced a **rising churn rate**, indicating that more customers are stopping their shopping activities on the platform.

By identifying churn patterns, Shopshop can make data-driven decisions to **strengthen loyalty** and **sustain long-term growth**.



Problems

➤ Churn Segment Identification

Customer **churn rates are rising**, particularly within **certain customer segments** — the company needs to **identify** which segments are most affected.

➤ Key Issues

The company seeks to understand the **key factors driving** customer churn.

➤ Strategic Recommendation

The ultimate goal is to help management **develop data-driven strategies** to **reduce churn** and **strengthen customer loyalty**.

DATA UNDERSTANDING

DATA UNDERSTANDING

The dataset contains 5,630 customer records and 20 variables related to customer behavior and engagement on the Shopshop e-commerce platform.

CustomerID	PreferedOrderCat
Churn	SatisfactionScore
Tenure	MaritalStatus
PreferredLoginDevice	NumberOfAddress
CityTier	Complain
WarehouseToHome	OrderAmountHikeFromlastYear
PreferredPaymentMode	CouponUsed
Gender	OrderCount
HourSpendOnApp	DaySinceLastOrder
NumberOfDeviceRegistered	CashbackAmount

PRE- DESCRIPTIVE STATISTIC

The dataset includes both categorical and numerical features, making it suitable for exploratory data analysis (EDA), segmentation, and predictive modeling to uncover churn behavior.

- **Unique Identifiers** : Each customer has a unique *CustomerID*.
- **Behavioral Attributes** : Variables such as *PreferredLoginDevice*, *PreferredPaymentMode*, and *PreferredOrderCat* capture shopping habits and preferences.
- **Engagement Metrics** : *Tenure*, *HourSpendOnApp*, and *NumberOfDeviceRegistered* describe how long and how actively customers engage with the platform.
- **Transaction Details** : Fields like *OrderCount*, *OrderAmountHikeFromLastYear*, and *CashbackAmount* provide insights into purchase frequency and spending patterns.
- **Satisfaction & Retention Factors**: *SatisfactionScore*, *Complain*, and *CouponUsed* offer signals related to customer experience and loyalty.
- **Target Variable**: Churn indicates whether a customer has stopped purchasing on the platform (1 = churned, 0 = active).

PRE-DESCRIPTIVE STATISTIC

General Insights

- The dataset contains **5,630 observations**, with no major missing values in core behavioral variables.
- The average **churn rate is 16.8%**, meaning **roughly 1 in 6 customers stopped shopping** on the platform.
- The average tenure is about **10.2 months, but varies widely (0–61)**, suggesting **customers have very different engagement lifespans**.

Experience & Satisfaction

- The mean satisfaction score is **3.1 out of 5**, with noticeable variation, suggesting room for service improvement.
- Complaints are low (0.28), but their correlation with churn can be further explored.

Behavior & Engagement

- Customers spend an **average of 2.9 hours** on the app, indicating moderate engagement.
- Most users register **around 3–4 devices**, showing multi-device behavior is common.
- The median order count is **3 orders per customer**, with a maximum of 16, showing a wide purchasing gap among users.

Transaction & Retention Metrics

- The average cashback amount is **177 ranging up to 325**, implying **diverse spending or promotional activity levels**.
- Order amount hike from last year **averages 15.7**, showing **gradual but steady customer growth** in order value.
- Coupons used are relatively **low on average (1.75 per customer)**, which may **indicate limited promotional engagement**.

HANDLING MISSING VALUES

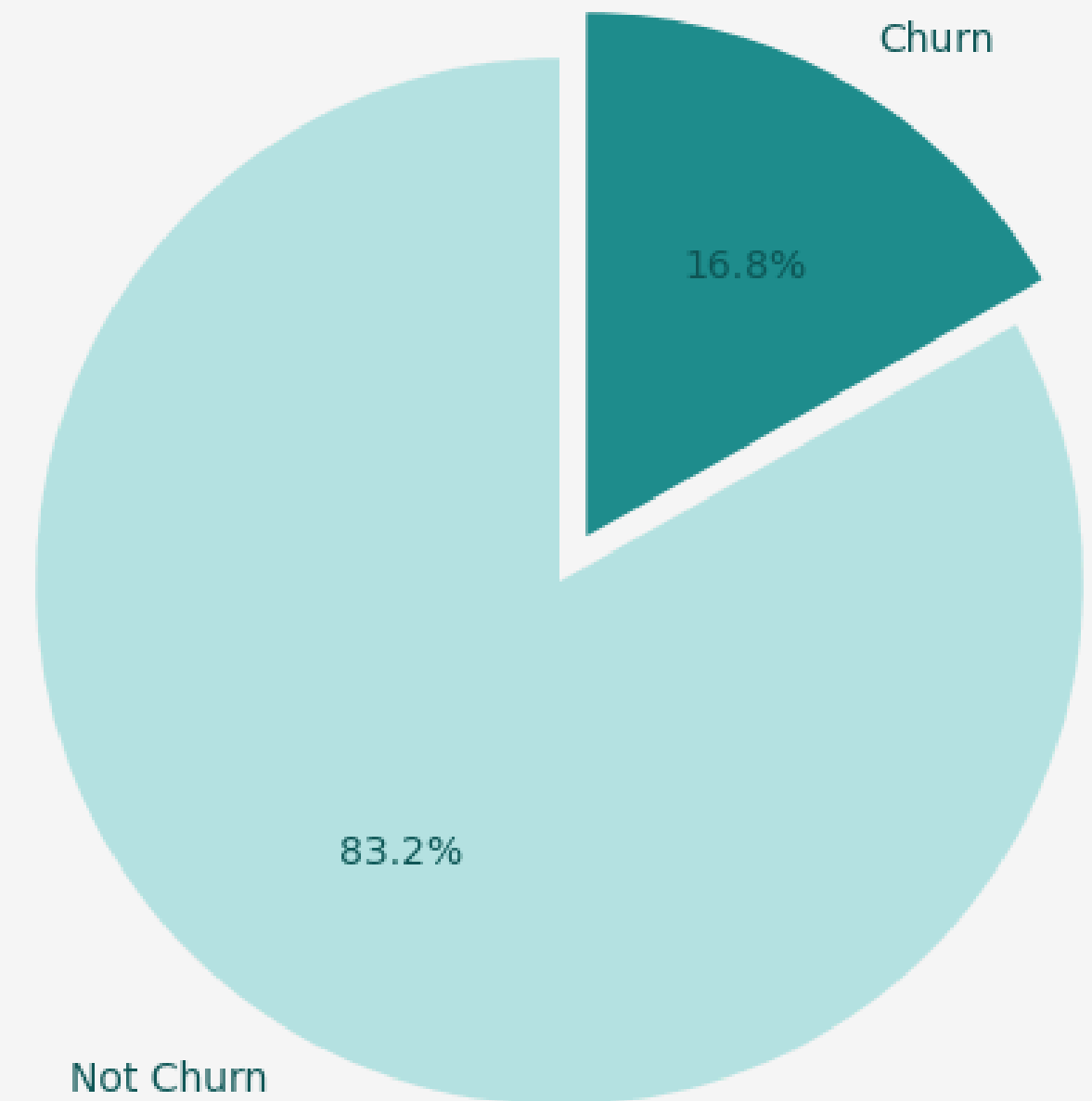
Column	Type	Missing Count	Handling Method	Rationale
Tenure	Numerical	264	Filled with median value	Tenure represents customer duration; median imputation minimizes outlier influence.
WarehouseToHome	Numerical	251	Filled with median value	Reflects delivery distance; median ensures a balanced central tendency without distortion.
HourSpendOnApp	Numerical	255	Filled with median value	Represents user engagement; median avoids bias from highly active users.
OrderAmountHikeFromLastYear	Numerical	265	Filled with median value	Indicates year-over-year order growth; median preserves realistic distribution.
CouponUsed	Numerical	256	Filled with median value	Counts of coupons used; median maintains neutrality and prevents inflation of low values.
OrderCount	Numerical	258	Filled with median value	Reflects customer purchasing frequency; median provides robust imputation for skewed data.
DaySinceLastOrder	Numerical	307	Left as missing (flagged or replaced with zero)	Missing values likely represent new or inactive users who haven't placed any order. Retaining or flagging helps capture this churn-related behavior.

EXPLORATORY DATA ANALYSIS

Churn vs non-churn

- **16.8% of customers have churned**, while 83.2% remain active.
- Indicates a **moderate churn level**, requiring stronger **customer retention strategies**.
- Next focus: **identify key churn drivers** to design data-driven actions — e.g., personalized promotions and improved user experience.

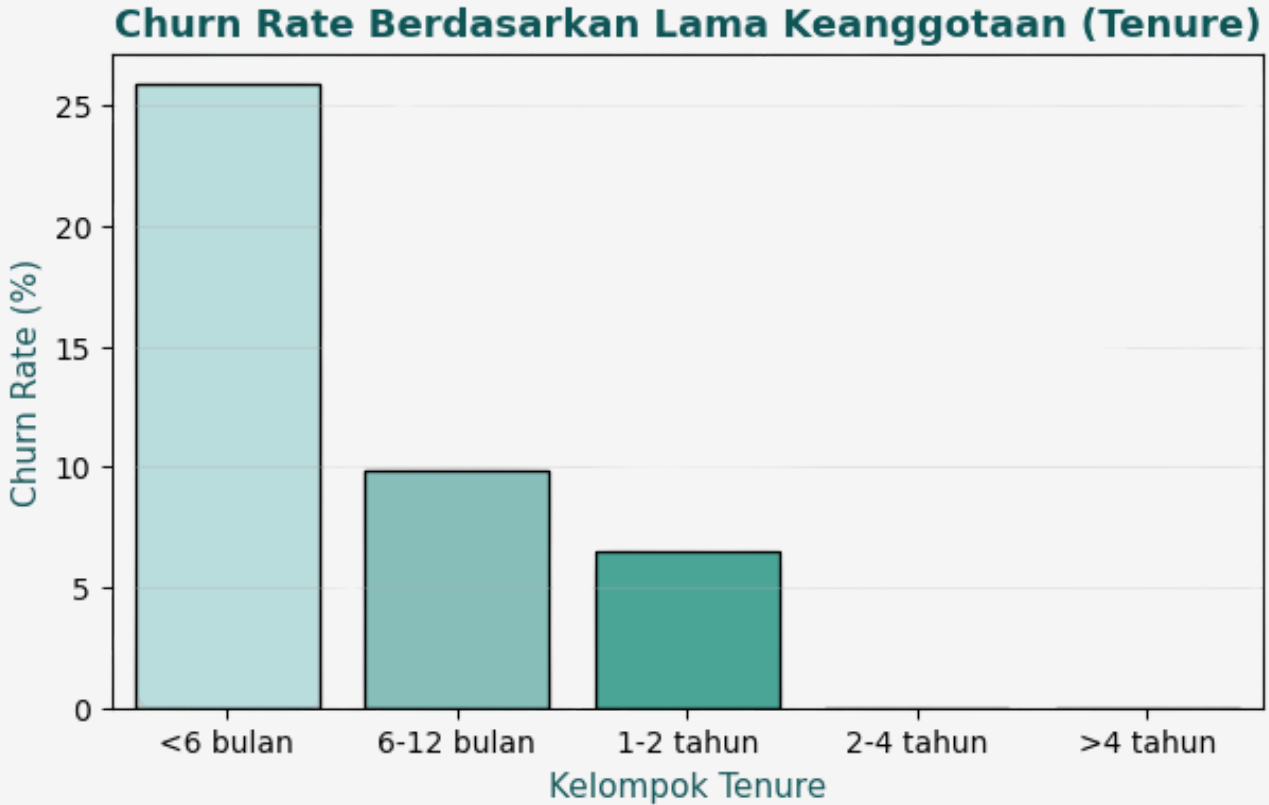
Proporsi Pelanggan Churn vs Tidak Churn



IDENTIFYING CHURN BASED ON CUSTOMER TENURE

- < 6 Months: Highest churn (25.9%) — nearly 1 in 4 new users leave early, showing there is a pattern **why people easily leave**.
- 6–12 Months: Churn drops to 9.9% — users start adapting but seasonal churn still occurs.
- 1–2 Years: Churn falls to 6.5% — customers show growing loyalty.
- > 2 Years: 0% churn — long-term users are highly engaged and act as brand advocates.

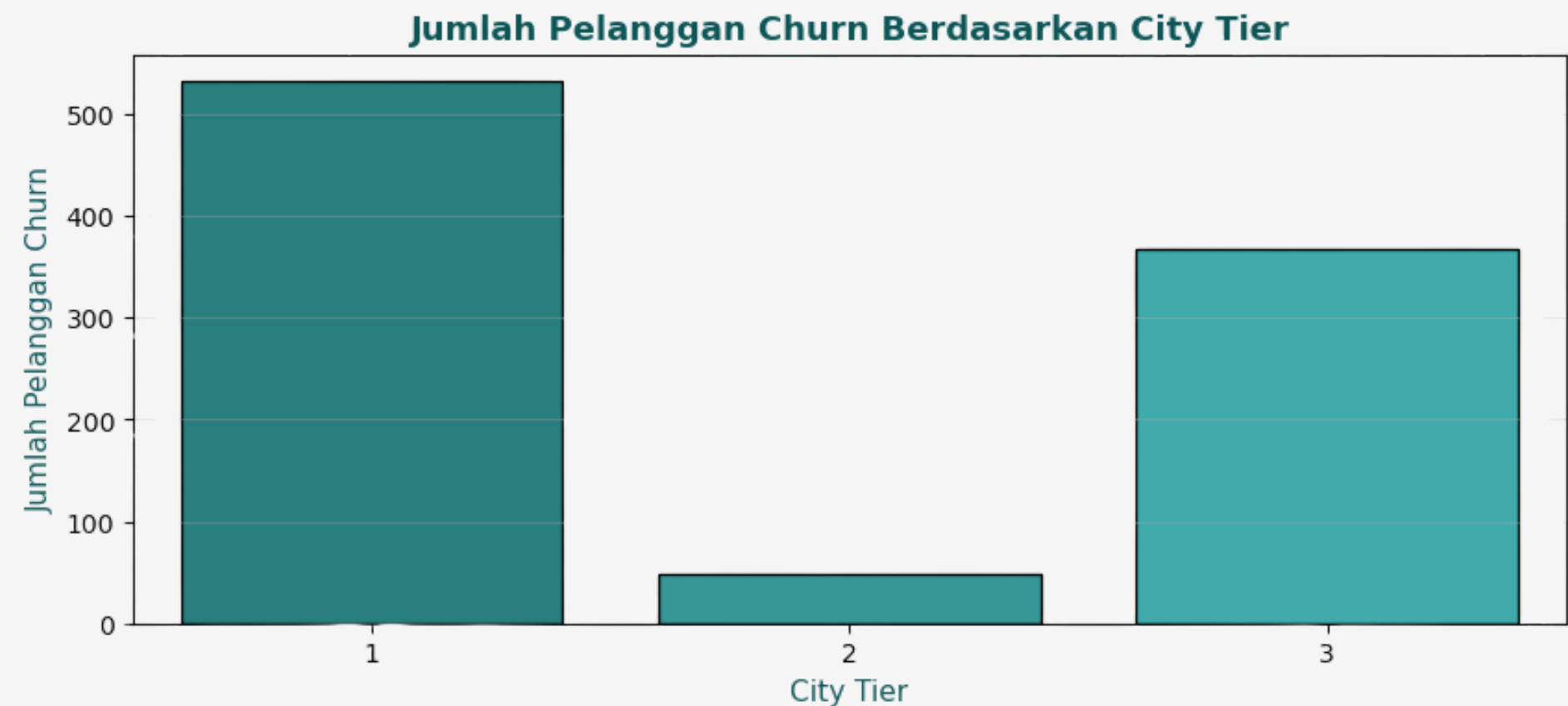
Tenure_Group	Total_Customers	Churn_Count	Churn_Rate
<6 bulan	1642	425	25.88
6-12 bulan	1584	156	9.85
1-2 tahun	1467	95	6.48
2-4 tahun	425	0	0
>4 tahun	4	0	0



IDENTIFYING CHURN BASED ON CITY TIER SEGMENT

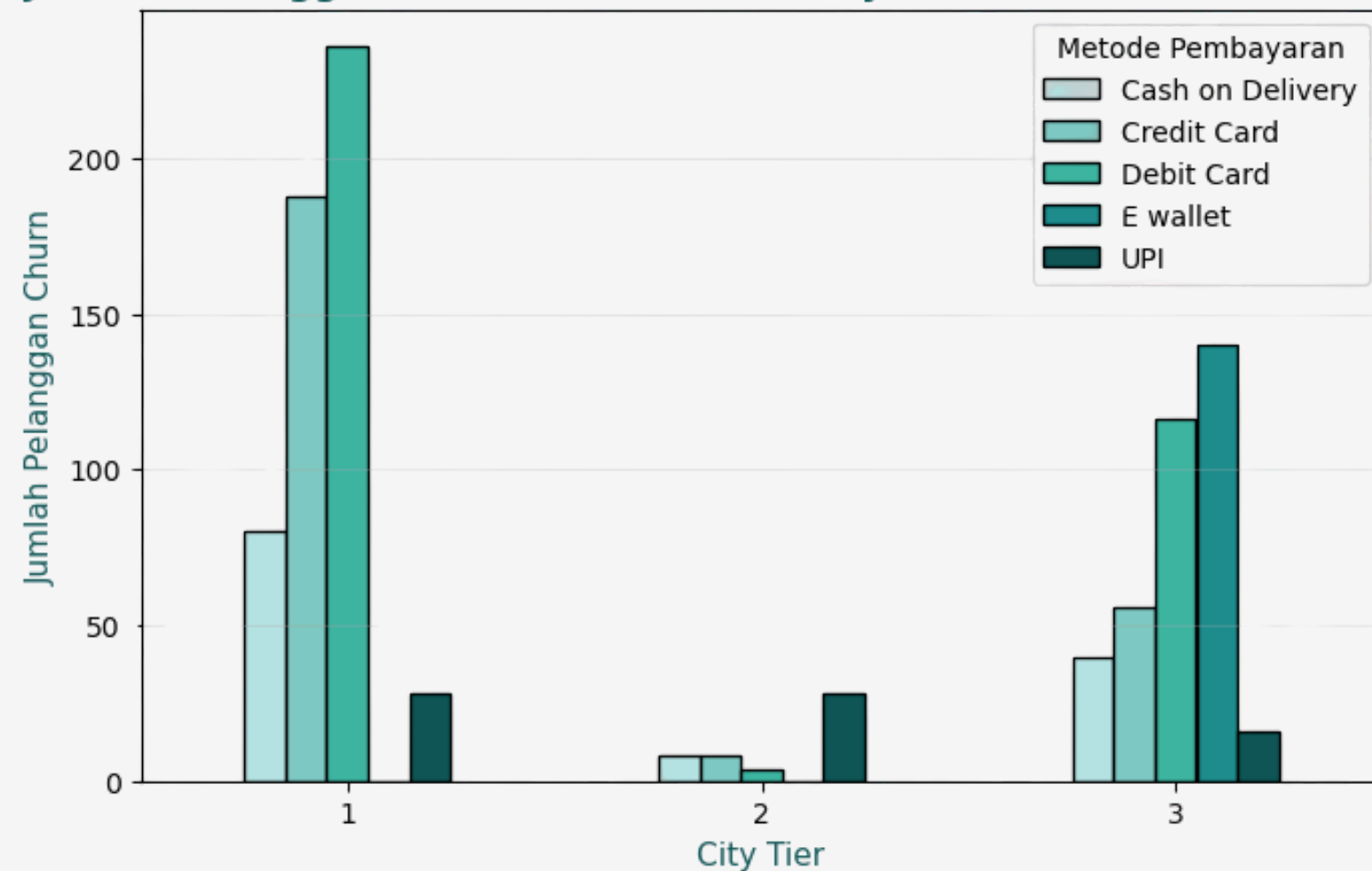
City Tier 1 shows the **highest churn volume (532 customers)** but the **lowest churn rate (14.5%)**, indicating that large cities have a **broad customer base that dilutes the churn proportion**. In contrast, City Tier 2 records a **lower churn volume (48)** yet a **higher churn rate (19.8%)**, reflecting **moderate customer engagement**. Meanwhile, **City Tier 3 has the highest churn rate (21.4%)**, making small-city customers **the most vulnerable segment**, with nearly one in five stopping their purchases — **signaling a need for stronger retention efforts**.

CityTier	Total_Customers	Churn_Count	Churn_Rate
1	3666	532	14.5%
2	242	48	19.8%
3	1722	368	21.4%



IDENTIFYING CHURN BASED ON CITY TIER SEGMENT

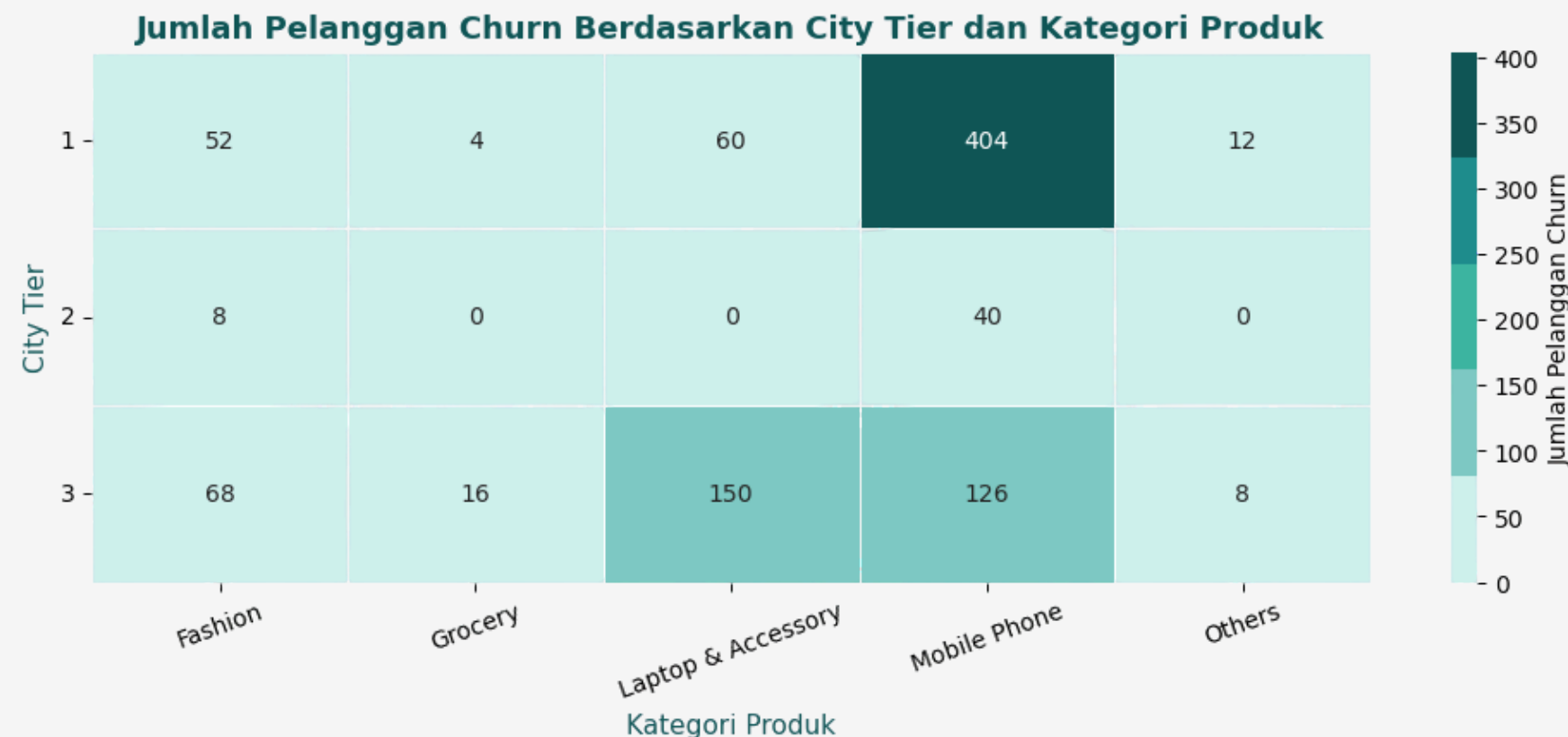
Jumlah Pelanggan Churn Berdasarkan City Tier dan Metode Pembayaran



CityTier x PreferredPaymentMode

- Tier 1 (Big Cities): Highest churn from **Debit** (236) & **Credit Card** (188) users
- Tier 2 (Mid Cities): Low, balanced churn — small, stable market; lower priority.
- Tier 3 (Small Cities): Churn spike in **E-Wallet** (140), **Debit** (116) and **Credit Card** (56)— new, less-loyal users drawn by rival promos.

IDENTIFYING CHURN BASED ON CITY TIER SEGMENT



City Tier x PreferOrderCat

Tier 1 – Big Cities:

- Highest churn in **Mobile Phone** (404 = 76%).
- Driven by long repurchase cycle and intense e-commerce price competition.
- Other categories (e.g., Laptop 60) show moderate churn impact.

Tier 2 – Mid Cities:

- Overall churn low (48) but still dominated by **Mobile Phone** (40).
- Consistent pattern → high-value product buyers churn after single purchase.

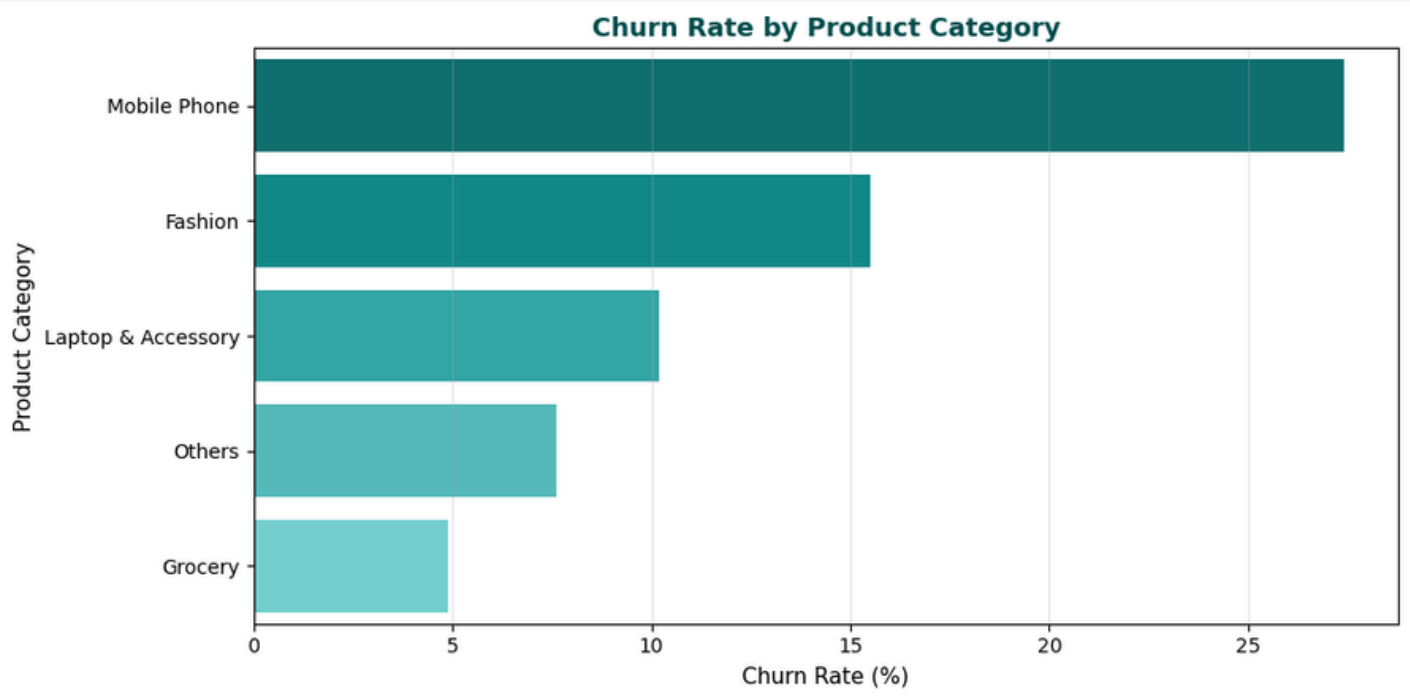
Tier 3 – Small Cities:

- High churn in **Laptop & Accessory** (150) and **Mobile Phone** (126).
- Purchases of expensive items rarely repeated → greater churn risk.

IDENTIFYING CHURN BASED ON CATEGORICAL SEGMENT

This analysis **identifies which product categories** have **the highest customer churn rate**. By examining churn distribution across product types, the company **can prioritize retention strategies for the most vulnerable segments**.

IDENTIFYING CHURN BASED ON CATEGORICAL SEGMENT



High-Risk Segments

- *Mobile Phone* – 27.4% Churn (570/2080 customers)
 - Highest churn and largest customer base.
 - Indications likely causes: **low repeat purchase frequency** and **price-driven switching to competitors**.
- *Fashion* – 15.5% Churn (128/826 customers)
 - Moderate , but churn-prone.
 - Indications likely causes: Customers **easily influenced by discounts, promotions, and seasonal trends**.

Moderate Segment

- *Laptop & Accessory* – 10.2% Churn (210/2050 customers)
 - Borderline churn level (slightly above safe zone).
 - Similar purchase pattern to Mobile Phones — **infrequent but high value**.

Low-Risk / Loyal Segments

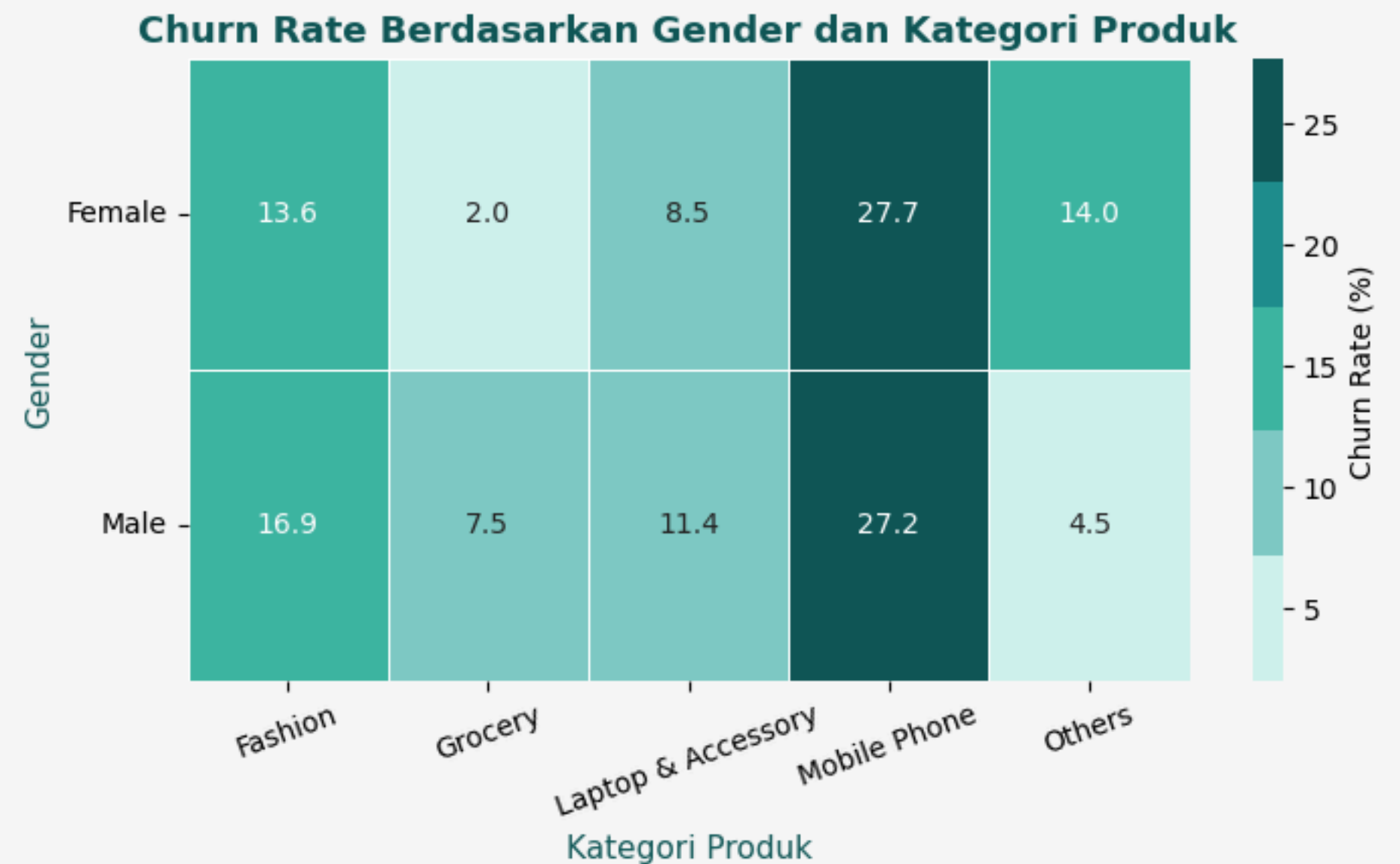
- *Others* – 7.6% and *Grocery* – 4.9%
 - Lowest churn; customers show **consistent buying behavior**.
 - Likely driven by **routine** or **necessity-based purchases**.

PreferedOrderCat	Total_Customers	Churn_Count	Churn_Rate
Mobile Phone	2080	570	27.4%
Fashion	826	128	15.5%
Laptop & Accessory	2050	210	10.2%
Others	264	20	7.6%
Grocery	410	20	4.9%

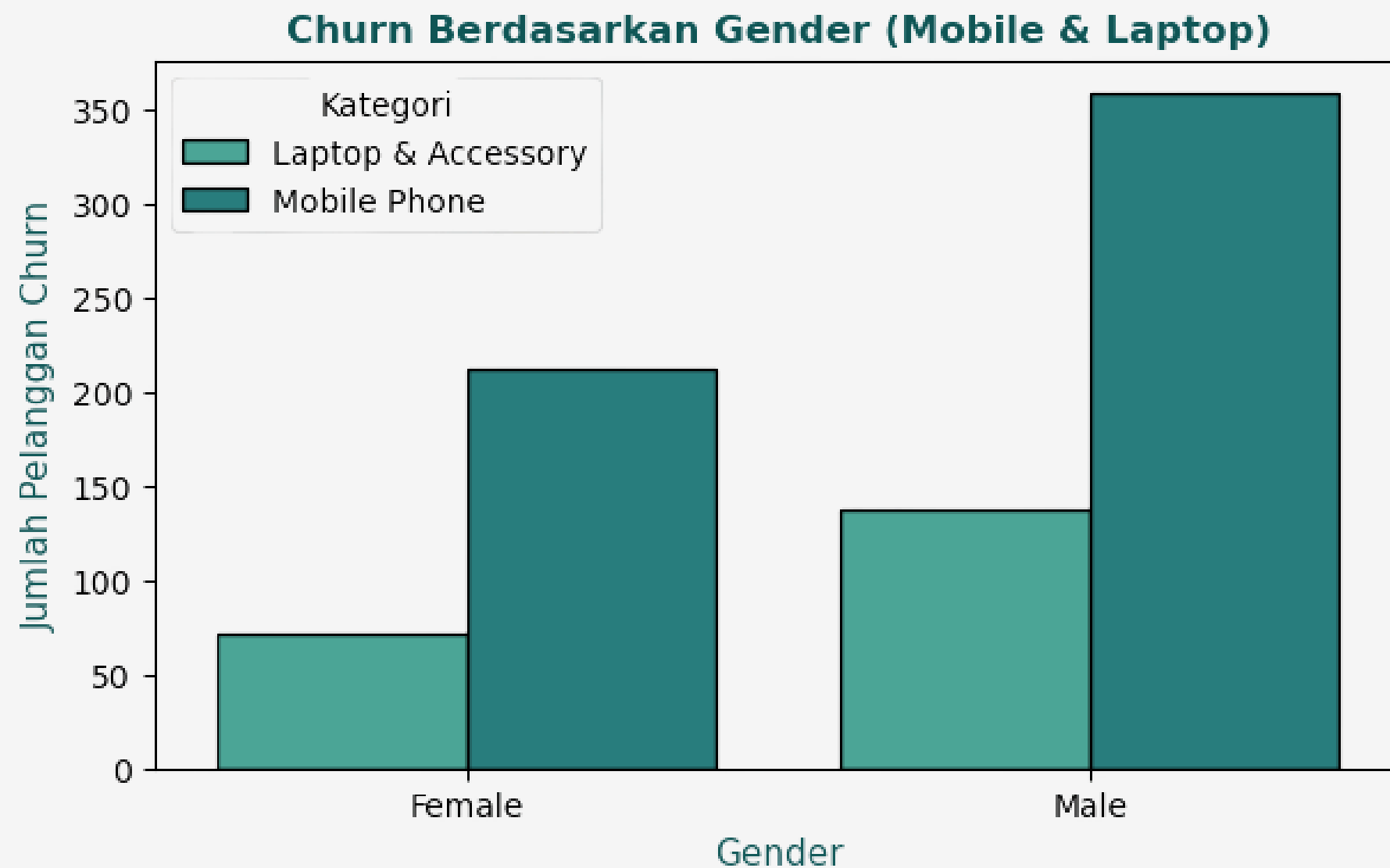
IDENTIFYING CHURN: GENDER BASED ON CATEGORICAL SEGMENT

Churn analysis by gender and product category shows that **Mobile Phone** has the **highest churn** for both genders ($\approx 27\%$), while **Grocery** has the lowest (2–7%). Male customers churn more in **Fashion** (16.9%) and **Laptop** (11.4%), whereas female customers remain stable but show higher churn in **Mobile Phone** (27.7%) and **Others** (14%).

This suggests that churn behavior **varies** by product type and gender preference.



IDENTIFYING CHURN: GENDER BASED ON CATEGORICAL SEGMENT (GADGET)



Categories like **Mobile Phone** show the highest overall churn across both genders. **Laptop & Accessory** churn is mainly driven by male customers, though at a smaller volume. This occurs because these products have **low purchase frequency**, and **churn often reflects one-time purchases** rather than **dissatisfaction**.

CONCLUSION & RECOMMENDATION

CONCLUSION

Churn is concentrated in **16.8% of total customers**, most of whom have a **tenure of less than six months** using the Shopshop platform.

CONCLUSION

Among these churned users, a significant share comes from **City Tier 1** and **Tier 3**, with **Debit Card** and **Credit Card** being the most common payment methods.

E-Wallet transactions in **Tier 3** also show a noticeable churn pattern, indicating that **digital payment users** in **smaller cities are still developing trust and loyalty** toward the platform. This suggests the need for **localized engagement programs** and **targeted incentives** to **strengthen customer retention** in these segments.

CONCLUSION

The *Mobile Phone* and *Laptop & Accessories* categories **contribute the most to the churn rate** (for **both genders**). From a consumer behavior perspective, customers in these categories tend to have **low purchase frequency** due to the **long product lifecycle** and **low-maintenance nature of gadgets**. Consequently, their churn behavior is often driven by **one-time purchase patterns** rather than **dissatisfaction**.

RECOMMENDATION

Target	Recommendation	Objective
Male (Mobile Phone)	Create a loyalty upgrade program (e.g., trade-in + bonus points).	Reduce churn among frequent device upgraders.
Male (Laptop)	Offer accessory bundles or extended warranty after purchase.	Improve retention in long purchase-cycle categories.
Female (Mobile Phone & Laptop)	Promote feminine or stylish gadget accessories through SKU push logic.	Attract emotional-driven mechanism female customers.
All Segments	Use AI-based personalized recommendations based on device history.	Increase promotional relevance and engagement .

RECOMMENDATION

- Establish **strategic partnerships with banks** to **offer promotional benefits** for **Debit and Credit Card** transactions.
- Launch **exclusive promotions** for customers using **Debit and Credit Card payments** (Discount, cashback, installment program, etc).
- Introduce **low-interest or 0% installment programs** to **boost interest** in **gadget-related categories** (mobile phones, laptops, and accessories).
- Provide special regional promotions in Tier 3 cities, focusing on **E-Wallet payment methods** to **strengthen engagement** and **loyalty**.